PCT

Report No. 85

UNITED STATES
DEPARTMENT OF THE INTERIOR
BUREAU OF MINES
HELIUM ACTIVITY
HELIUM RESEARCH CENTER

INTERNAL REPORT

NON-LINEAR REGRESSION AND THE PRINCIPLE OF LEAST SQUARES

THE METHOD OF EVALUATING THE CONSTANTS AND THE CALCULATION OF VARIANCES AND COVARIANCES

BY

Robert E. Barieau

B. J. Dalton

BRANCH Fundamental Research

PROJECT NO. 4335

DATE April 1966

AMARILLO, TEXAS

Report No. 85

HELIUM RESEARCH CENTER

INTERNAL REPORT

NON-LINEAR REGRESSION AND THE PRINCIPLE OF LEAST SQUARES. THE METHOD OF EVALUATING THE CONSTANTS AND THE CALCULATION OF VARIANCES AND COVARIANCES

By

Robert E. Barieau and B. J. Dalton

Branch of Fundamental Research

Project 4335

April 1966

CONTENTS

	Page
Abstract	3
Introduction	3
Evaluation of the constants	7
Calculation of variances and covariances	33
References	55

NON-LINEAR REGRESSION AND THE PRINCIPLE OF LEAST SQUARES. THE METHOD OF EVALUATING THE CONSTANTS AND THE CALCULATION OF VARIANCES AND COVARIANCES.

by

Robert E. Barieau[1/] and B. J. Dalton[2/]

ABSTRACT

This report is concerned with the necessary mathematical equations in order to accomplish the following objectives: 1. to evaluate the parameters such that the sum of the weighted squares of the residuals of an experimental observable be a true minimum, regardless of the functional relationship between the variables and these parameters; 2. to evaluate all variances and covariances of the parameters evaluated by means of the law for the propagation of errors; and 3. to insure that a true minimum will always be obtained. The equations presented in this report were developed for a three-parameter problem.

INTRODUCTION

The Helium Research Center, Bureau of Mines, has as one of its long-range objectives the development of an equation of state for

---

1/ Supervisory Research Chemist, Project Leader, Thermodynamics, Helium Research Center, Bureau of Mines, Amarillo, Texas.

2/ Research Chemist, Helium Research Center, Bureau of Mines, Amarillo, Texas.

Work on manuscript completed December 1965.

helium that will reproduce the data to within the accuracy with which the data are known.

The Helium Research Center also has an experimental program for obtaining PVT data on gases and mixtures by the Burnett (2)[3/] method.

3/ Underlined numbers in parentheses refer to items in the list of references at the end of this report.

In the Burnett method, one of the constants that must be evaluated is the volume-ratio of two containers. By the theory of the method, this constant is inherently non-linear. It was therefore decided to develop a capability for handling non-linear regression problems. This report gives the principles of the method the Helium Research Center uses in such problems.

Our method differs in several important respects from methods currently in use. In solving non-linear problems, the problem must be linearized and an iteration technique used to obtain the solution. All texts on non-linear regression, of which we are aware, linearize the problem before the normal equations are formed. This method is known as the Gauss-Newton method. In our method, the exact normal equations are formed, and the problem is linearized by expanding the exact normal equations in a Taylor's series expansion retaining the first two terms. This method is known as the Newton-Raphson method (7). The only work that we have been able to locate that uses this method in non-linear least squares problems is that of Strand, Kohl, and Bonham (8).

These two different methods, as will be shown later, lead to the same least squares solution provided the iteration procedure converges to an answer. We have found that if one starts within the region of convergence, the Newton-Raphson method converges more rapidly than the Gauss-Newton method. This is one advantage of the method we have chosen.

If on applying the Gauss-Newton or the Newton-Raphson method, the problem is diverging after the first iteration, a method must be found that will lead to convergence. One of the methods that may be tried at this stage is the negative gradient or method of steepest descent. If the step in the direction of the negative gradient is small enough, this method must lead to a smaller sum of the squares of the deviations. The problem with this method in the past has been deciding on the size of the step to be taken. If the Newton-Raphson method has been used in developing the normal equations for the first iteration, then the size of the step to be taken in the direction of the negative gradient can be evaluated very simply from the coefficients appearing in the normal equations. This is the second advantage of the method we have chosen.

The third advantage involves the calculation of the variances and covariances of the constants evaluated. As far as we are aware, all authors and all programs available calculate variances and covariances on the assumption that the formulas that apply to linear problems will apply to non-linear problems once the non-linear problems have been linearized. We reject this assumption, preferring to calculate

variances and covariances from the fundamental definition of these quantities and the law for the propagation of errors. (1, 5).

Some authors (6) claim that the least squares values of the constants evaluated in a non-linear problem are biased and should be corrected. All the proofs of this, that we have seen, assume the deviations are distributed with zero mean. This, of course, is never true in a non-linear problem unless this condition is imposed as a constraint. Further, the principle of least squares maximizes the probability that the deviations are equal to the true random errors. This is true for both linear and non-linear problems. This being true, we fail to see how any solution can be better than the least squares solution. We therefore take the least squares solution as being non-biased and apply no correction.

We have set the following objectives for our method.

1. To evaluate the parameters so that the sum of the weighted squares of the residuals of an experimental observable is a true minimum.

2. Objective 1 is to be accomplished even though the functional relationship between the observables and parameters is such that the observable involved in the minimum of the sum of the weighted squares of the residuals cannot be explicitly expressed as a function of other observables and the parameters.

3. All variances and covariances are to be calculated, with no approximations, by means of the law for the propagation of errors.

4. The method is to be such that a true minimum will always be obtained.

This report is concerned with the necessary mathematical equations in order to accomplish the above-named objectives. The equations are developed for a three-parameter problem. The extension to the evaluation of more parameters should be obvious.

EVALUATION OF THE CONSTANTS

Suppose one has experimentally determined a set of n data points x_i, y_i. Let the functional relationship between x and y and the parameters A, B, and C be given by

$$F(y,x,A,B,C) = 0 \tag{1}$$

We assume that there are no random errors in the x_i's and that random errors occur in the observed y_i's.

Now suppose we have evaluated the constants A, B, and C by some means or other. Then, because of random errors in y_i, equation (1) will not be exactly satisfied when the observed y_i and x_i values are substituted in equation (1). We will let F_i be the numerical value of F, when the observed y_i and x_i values are substituted in equation (1). Thus,

$$F_i = F(y_{i(o)},x_i,A,B,C) \tag{2}$$

Now when x_i is substituted in equation (1), we may solve for y_i so that equation (1) is satisfied exactly. We will designate this y_i as $y_{i(calc)}$. Thus,

$$F(y_{i(calc)},x_i,A,B,C) = 0 \tag{3}$$

The residual of y_i is given by

$$Y_i = y_{i(o)} - y_{i(calc)} \tag{4}$$

Now Y_i, the residual of y_i, is the difference between the observed and calculated values. This is not the true random error in our observed y_i because we do not know the true value of y_i. However, we can maximize the probability that our Y_i's are equal to the true random errors, and this is just what the principle of least squares does. The principle of least squares says that we maximize the probability that the Y_i's represent the true random errors by minimizing the sum of the weighted squares of the residuals.

Thus, the function to be minimized is given by

$$R = \sum_{i=1}^{n} w_{y_{i(o)}} Y_i^2 \tag{5}$$

where $w_{y_{i(o)}}$ is the weight assigned to $y_{i(o)}$. R is a function of the constants to be evaluated: A, B, and C. The condition that R be a minimum is determined by

$$\frac{1}{2}\left(\frac{\partial R}{\partial A}\right)_{B,C} = \sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left(\frac{\partial Y_i}{\partial A}\right) = 0 \tag{6}$$

$$\frac{1}{2}\left(\frac{\partial R}{\partial B}\right)_{A,C} = \sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left(\frac{\partial Y_i}{\partial B}\right) = 0 \tag{7}$$

and

$$\frac{1}{2}\left(\frac{\partial R}{\partial C}\right)_{A,B} = \sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left(\frac{\partial Y_i}{\partial C}\right) = 0 \tag{8}$$

In the application of equations (6), (7), and (8), the observed y_i's and x_i's are to be held constant in the derivatives: $\left(\frac{\partial R}{\partial A}\right)_{B,C}$, $\left(\frac{\partial R}{\partial B}\right)_{A,C}$, $\left(\frac{\partial R}{\partial C}\right)_{A,B}$. Equations (6), (7), and (8) are the exact normal equations.

If Y_i is non-linear in the undetermined constants, then the solutions of our normal equations will not be straightforward. It will be necessary to solve them by an iterative technique in which values of the constants are assumed. This is done by expanding Y_i, $\left(\frac{\partial Y_i}{\partial A}\right)$, $\left(\frac{\partial Y_i}{\partial B}\right)$, and $\left(\frac{\partial Y_i}{\partial C}\right)$ in a Taylor's series expansion about an approximate solution, Y_i^o, retaining only the first two terms. Thus,

$$Y_i = Y_i^o + \left(\frac{\partial Y_i}{\partial A}\right)^o \Delta A + \left(\frac{\partial Y_i}{\partial B}\right)^o \Delta B + \left(\frac{\partial Y_i}{\partial C}\right)^o \Delta C \tag{9}$$

$$\left(\frac{\partial Y_i}{\partial A}\right) = \left(\frac{\partial Y_i}{\partial A}\right)^o + \left(\frac{\partial^2 Y_i}{\partial A^2}\right)^o \Delta A + \left(\frac{\partial^2 Y_i}{\partial A \partial B}\right)^o \Delta B + \left(\frac{\partial^2 Y_i}{\partial A \partial C}\right)^o \Delta C \tag{10}$$

$$\left(\frac{\partial Y_i}{\partial B}\right) = \left(\frac{\partial Y_i}{\partial B}\right)^o + \left(\frac{\partial^2 Y_i}{\partial B \partial A}\right)^o \Delta A + \left(\frac{\partial^2 Y_i}{\partial B^2}\right)^o \Delta B + \left(\frac{\partial^2 Y_i}{\partial B \partial C}\right)^o \Delta C \tag{11}$$

$$\left(\frac{\partial Y_i}{\partial C}\right) = \left(\frac{\partial Y_i}{\partial C}\right)^o + \left(\frac{\partial^2 Y_i}{\partial C \partial A}\right)^o \Delta A + \left(\frac{\partial^2 Y_i}{\partial C \partial B}\right)^o \Delta B + \left(\frac{\partial^2 Y_i}{\partial C^2}\right)^o \Delta C \tag{12}$$

where the quantities ΔA, ΔB, and ΔC are defined as

$$\Delta A = A - A_o$$

$$\Delta B = B - B_o$$

$$\Delta C = C - C_o$$

where A, B, and C are our undetermined constants, and A_o, B_o, and C_o are approximate values for these quantities. Then to first order in the Δ's

$$Y_i\left(\frac{\partial Y_i}{\partial A}\right) = \left[Y_i^o\left(\frac{\partial Y_i}{\partial A}\right)^o + \left[\left(\frac{\partial Y_i}{\partial A}\right)^{o^2} + Y_i^o\left(\frac{\partial^2 Y_i}{\partial A^2}\right)^o\right]\Delta A + \left[\left(\frac{\partial Y_i}{\partial B}\right)^o\left(\frac{\partial Y_i}{\partial A}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial B\partial A}\right)^o\right]\Delta B + \left[\left(\frac{\partial Y_i}{\partial C}\right)^o\left(\frac{\partial Y_i}{\partial A}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial C\partial A}\right)^o\right]\Delta C \right] \tag{13}$$

$$Y_i\left(\frac{\partial Y_i}{\partial B}\right) = \left[Y_i^o\left(\frac{\partial Y_i}{\partial B}\right)^o + \left[\left(\frac{\partial Y_i}{\partial A}\right)^o\left(\frac{\partial Y_i}{\partial B}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial A\partial B}\right)^o\right]\Delta A + \left[\left(\frac{\partial Y_i}{\partial B}\right)^{o^2} + Y_i^o\left(\frac{\partial^2 Y_i}{\partial B^2}\right)^o\right]\Delta B + \left[\left(\frac{\partial Y_i}{\partial B}\right)^o\left(\frac{\partial Y_i}{\partial C}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial B\partial C}\right)^o\right]\Delta C \right] \tag{14}$$

$$Y_i\left(\frac{\partial Y_i}{\partial C}\right) = \left[Y_i^o\left(\frac{\partial Y_i}{\partial C}\right)^o + \left[\left(\frac{\partial Y_i}{\partial A}\right)^o\left(\frac{\partial Y_i}{\partial C}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial A\partial C}\right)^o\right]\Delta A + \left[\left(\frac{\partial Y_i}{\partial B}\right)^o\left(\frac{\partial Y_i}{\partial C}\right)^o + Y_i^o\left(\frac{\partial^2 Y_i}{\partial B\partial C}\right)^o\right]\Delta B + \left[\left(\frac{\partial Y_i}{\partial C}\right)^{o^2} + Y_i^o\left(\frac{\partial^2 Y_i}{\partial C^2}\right)^o\right]\Delta C \right] \tag{15}$$

Substituting equations (13), (14), and (15) into our normal equations (6), (7), and (8), we have

$$a_1 \Delta A + b_1 \Delta B + c_1 \Delta C = m_1 \tag{16}$$

$$a_2 \Delta A + b_2 \Delta B + c_2 \Delta C = m_2 \tag{17}$$

$$a_3 \Delta A + b_3 \Delta B + c_3 \Delta C = m_3 \tag{18}$$

where

$$a_1 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial A}\right)^{o^2} + Y_i^o \left(\frac{\partial^2 Y_i}{\partial A^2}\right)^o \right] \tag{19}$$

$$a_2 = b_1 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial B}\right)^o \left(\frac{\partial Y_i}{\partial A}\right)^o + Y_i^o \left(\frac{\partial^2 Y_i}{\partial A \partial B}\right)^o \right] \tag{20}$$

$$a_3 = c_1 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial A}\right)^o \left(\frac{\partial Y_i}{\partial C}\right)^o + Y_i^o \left(\frac{\partial^2 Y_i}{\partial A \partial C}\right)^o \right] \tag{21}$$

$$b_2 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial B}\right)^{o^2} + Y_i^o \left(\frac{\partial^2 Y_i}{\partial B^2}\right)^o \right] \tag{22}$$

$$b_3 = c_2 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial B}\right)^o \left(\frac{\partial Y_i}{\partial C}\right)^o + Y_i^o \left(\frac{\partial^2 Y_i}{\partial B \partial C}\right)^o \right] \tag{23}$$

$$c_3 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial C}\right)^{o^2} + Y_i^o \left(\frac{\partial^2 Y_i}{\partial C^2}\right)^o \right] \tag{24}$$

$$m_1 = -\sum_{i=1}^{n} w_{y_{i(o)}} Y_i^o \left(\frac{\partial Y_i}{\partial A}\right)^o \quad (25)$$

$$m_2 = -\sum_{i=1}^{n} w_{y_{i(o)}} Y_i^o \left(\frac{\partial Y_i}{\partial B}\right)^o \quad (26)$$

$$m_3 = -\sum_{i=1}^{n} w_{y_{i(o)}} Y_i^o \left(\frac{\partial Y_i}{\partial C}\right)^o \quad (27)$$

The solutions to equations (16), (17), and (18) are

$$D_o \Delta A = D_1 m_1 + D_2 m_2 + D_3 m_3 \quad (28)$$

$$D_o \Delta B = D_4 m_1 + D_5 m_2 + D_6 m_3 \quad (29)$$

$$D_o \Delta C = D_7 m_1 + D_8 m_2 + D_9 m_3 \quad (30)$$

where

$$D_1 = b_2 c_3 - b_3 c_2 \quad (31)$$

$$D_2 = b_3 c_1 - b_1 c_3 \quad (32)$$

$$D_3 = b_1 c_2 - b_2 c_1 \quad (33)$$

$$D_4 = a_3 c_2 - a_2 c_3 \quad (34)$$

$$D_5 = a_1 c_3 - a_3 c_1 \quad (35)$$

$$D_6 = a_2 c_1 - a_1 c_2 \quad (36)$$

$$D_7 = a_2 b_3 - a_3 b_2 \quad (37)$$

$$D_8 = a_3 b_1 - a_1 b_3 \quad (38)$$

$$D_9 = a_1 b_2 - a_2 b_1 \quad (39)$$

$$D_o = a_1 D_1 + a_2 D_2 + a_3 D_3 \tag{40}$$

The solutions of equations (28), (29), and (30) give the corrections to be applied to the assumed values of our undetermined constants.

In the Gauss-Newton method of linearization, the second term in the summations of the a's, b's, and c's of equations (19)-(24) is neglected; this does not lead to an error as long as the method converges because the exact solutions of equations (16), (17), and (18) are: $\Delta A = \Delta B = \Delta C = 0$.

When the functional relationship between the observables is such that $y_{i(calc)}$ cannot be solved for explicitly, it will be necessary to solve equation (3) for $y_{i(calc)}$ by a series of approximations. Let us expand F in a truncated Taylor's series expansion about the point x_i, $y_{i(o)}$:

$$F = F_i + \left(\frac{\partial F}{\partial y}\right)_{\substack{x,A,B,C\\ y=y_{i(o)}\\ x=x_i}} \left(y_{i(calc)} - y_{i(o)}\right) = 0 \tag{41}$$

or

$$y_{i(o)} - y_{i(calc)} = \frac{F_i}{\left(\frac{\partial F}{\partial y}\right)_{\substack{x,A,B,C\\ y=y_{i(o)}\\ x=x_i}}} \tag{42}$$

where in equations (41) and (42) the symbol $\left(\frac{\partial F}{\partial y}\right)_{\substack{x,A,B,C\\ y=y_{i(o)}\\ x=x_i}}$ means that

the derivative of F with regard to y, keeping x, A, B, and C constant, is to be evaluated at the point

$$y = y_{i\,(o)}$$

$$x = x_i$$

The solution of equation (42) gives the first approximation for $y_{i(calc)}$. We designate this value as $y_{i(calc)_1}$. This value is then substituted into equation (3), and if $y_{i(calc)_1}$ is not the exact answer, equation (3) will not be satisfied exactly. We designate this value of F as $F_{i(calc)_1}$. Thus,

$$F_{i\,(calc)_1} = F(x_i, y_{i(calc)_1}, A, B, C) \tag{43}$$

Then the second approximation, $y_{i(calc)_2}$, of $y_{i(calc)}$ is obtained from the expression

$$y_{i(calc)_1} - y_{i(calc)_2} = \frac{F_{i(calc)_1}}{\left(\frac{\partial F}{\partial y}\right)_{\substack{x,A,B,C\\ y=y_{i(calc)_1}\\ x=x_i}}} \tag{44}$$

This iteration is repeated until equation (3) is satisfied to within any amount we wish to specify.

Once we have $y_{i(calc)}$ and $y_{i(o)}$, then $Y_i = y_{i(o)} - y_{i(calc)}$, and Y_i^2 can be calculated. Then if $w_{y_{i(o)}}$ is known or has been assigned, R may be calculated by means of equation (4).

If the $y_{i(o)}$'s all have the same precision index, they will all have the same weight and $w_{y_{i(o)}} = 1$. If the $y_{i(o)}$'s do not all have the same precision index, then

$$w_{y_{i(o)}} = \frac{L^2}{S^2_{y_{i(o)}}} \tag{45}$$

where L is a constant and $S^2_{y_{i(o)}}$ is the variance of $y_{i(o)}$. In a particular problem, it may be necessary to assume that $w_{y_{i(o)}} = 1$ in the beginning. However, if this is done, the residuals, $\left[y_{i(o)} - y_{i(calc)}\right]$, should be examined to see if there is any statistical evidence for the residuals squared being a function of y. Any assumption as to the variance, $S^2_{y_{i(o)}}$, being a function of y can always be checked by examining the residuals. In any event, $w_{y_{i(o)}}$ is not a function of the constants to be evaluated.

We now proceed to develop the equations needed to calculate the coefficients of our normal equations. Differentiating equation (4) with regard to A, keeping x_i, $y_{i(o)}$, B, and C constant, we have

$$\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} = -\left(\frac{\partial y_{i(calc)}}{\partial A}\right)_{x_i, B, C} \tag{46}$$

Differentiating equation (3) with regard to A, keeping x_i, B, and C constant, we have

$$\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C} \left(\frac{\partial y_{i(calc)}}{\partial A}\right)_{x_i, B, C} + \left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C} = 0 \tag{47}$$

or

$$\left(\frac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C} = - \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \tag{48}$$

and if we substitute equation (48) in equation (46), we have

$$\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} = \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \tag{49}$$

Similarly, it can be shown that

$$\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} = \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \tag{50}$$

and that

$$\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} = \frac{\left(\frac{\partial F}{\partial C}\right)_{x_i,y_{i(calc)},A,B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \tag{51}$$

Differentiating equation (49) with regard to A, keeping x_i, $y_{i(o)}$, B, and C constant, we have .

$$\left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i,y_{i(o)},B,C} = \left[\begin{array}{c} \dfrac{\left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \\ \\ - \dfrac{\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} \end{array}\right] \tag{52}$$

Now

$$\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C} = G\left(y_{i(calc)},A,B,C\right)$$

$$\left[d\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,B,C} = \left[\begin{array}{l} \left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,y_{i(calc)},B,C} dA \\ \\ + \left[\dfrac{\partial}{\partial y_{i(calc)}}\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,A,B,C} dy_{i(calc)} \end{array}\right] \tag{53}$$

and

$$\left[\frac{\partial}{\partial A}\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,B,C} = \left[\begin{array}{l} \left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,y_{i(calc)},B,C} \\ \\ + \left[\dfrac{\partial}{\partial y_{i(calc)}}\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,A,B,C}\left(\dfrac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C} \end{array}\right] \tag{54}$$

Also,

$$\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C} = g\left(y_{i(calc)},A,B,C\right)$$

$$\left[d\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,B,C} = \left[\begin{array}{l} \left[\dfrac{\partial}{\partial y_{i(calc)}}\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,A,B,C} dy_{i(calc)} \\ \\ +\left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,y_{i(calc)},B,C} dA \end{array}\right] \quad (55)$$

and

$$\left[\frac{\partial}{\partial A}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,B,C} = \left[\begin{array}{l} \left[\dfrac{\partial}{\partial A}\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,y_{i(calc)},B,C} \\ \\ +\left[\dfrac{\partial}{\partial y_{i(calc)}}\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,A,B,C}\left(\dfrac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C} \end{array}\right] \quad (56)$$

When we substitute equations (54) and (56) into equation (52), we find

$$\left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i,y_{i(o)},B,C} = \left[\begin{array}{l} \dfrac{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\left(\dfrac{\partial^2 F}{\partial A^2}\right)_{x_i,y_{i(calc)},B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} \\[2ex] + \dfrac{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\left(\dfrac{\partial^2 F}{\partial y_{i(calc)}\partial A}\right)_{x_i,B,C}\left(\dfrac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} \\[2ex] - \dfrac{\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\left(\dfrac{\partial^2 F}{\partial A \partial y_{i(calc)}}\right)_{x_i,B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} \\[2ex] - \dfrac{\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\left(\dfrac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i,A,B,C}\left(\dfrac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} \end{array}\right] \tag{57}$$

But from equation (48)

$$\left(\frac{\partial y_{i(calc)}}{\partial A}\right)_{x_i,B,C} = -\frac{\left(\dfrac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}}{\left(\dfrac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \tag{48}$$

When we substitute equation (48) in equation (57), we see that

$$\left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i, y_{i(o)}, B, C} = \left[\frac{\left(\frac{\partial^2 F}{\partial A^2}\right)_{x_i, y_{i(calc)}, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{2\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\left(\frac{\partial^2 F}{\partial A \partial y_{i(calc)}}\right)_{x_i, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} + \frac{\left(\frac{\partial F}{\partial A}\right)^2_{x_i, y_{i(calc)}, B, C}\left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i, A, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i, A, B, C}}\right] \quad (58)$$

Similarly,

$$\left(\frac{\partial^2 Y_i}{\partial B^2}\right)_{x_i, y_{i(o)}, A, C} = \left[\frac{\left(\frac{\partial^2 F}{\partial B^2}\right)_{x_i, y_{i(calc)}, A, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{2\left(\frac{\partial F}{\partial B}\right)_{x_i, y_{i(calc)}, A, C}\left(\frac{\partial^2 F}{\partial B \partial y_{i(calc)}}\right)_{x_i, A, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} + \frac{\left(\frac{\partial F}{\partial B}\right)^2_{x_i, y_{i(calc)}, A, C}\left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i, A, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i, A, B, C}}\right] \quad (59)$$

and

$$\left(\frac{\partial^2 Y_i}{\partial C^2}\right)_{x_i, y_{i(o)}, A, B} = \left[\frac{\left(\frac{\partial^2 F}{\partial C^2}\right)_{x_i, y_{i(calc)}, A, B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{2\left(\frac{\partial F}{\partial C}\right)_{x_i, y_{i(calc)}, A, B}\left(\frac{\partial^2 F}{\partial C \partial y_{i(calc)}}\right)_{x_i, A, B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} + \frac{\left(\frac{\partial F}{\partial C}\right)^2_{x_i, y_{i(calc)}, A, B}\left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i, A, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i, A, B, C}}\right] \quad (60)$$

Differentiating equation (49) with regard to B, keeping $x_i, y_{i(o)}, A, C$ fixed, we have

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{i(o)}, A, C} = \left[\frac{\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\right]_{x_i, A, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}\right]_{x_i, A, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}}\right] \quad (61)$$

Let us differentiate $\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}$ with regard to B, keeping x_i, A, and C constant. When we do this, we get

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\right]_{x_i, A, C} = \left[\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\right]_{x_i, y_{i(calc)}, A, C} + \left[\frac{\partial}{\partial y_{i(calc)}}\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C}\right]_{x_i, A, B, C}\left(\frac{\partial y_{i(calc)}}{\partial B}\right)_{x_i, A, C}\right] \quad (62)$$

But

$$\left(\frac{\partial y_{i(calc)}}{\partial B}\right)_{x_i,A,C} = -\frac{\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} \quad (63)$$

and if we substitute equation (63) in equation (62), we see that

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\right]_{x_i,A,C} = \left[\left(\frac{\partial^2 F}{\partial A \partial B}\right)_{x_i,y_{i(calc)},C} - \left(\frac{\partial^2 F}{\partial A \partial y_{i(calc)}}\right)_{x_i,B,C} \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}}\right] \quad (64)$$

Now if we differentiate $\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}$ with regard to B, holding x_i, A, and C constant, we get

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,A,C} = \left[\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,y_{i(calc)},A,C} + \left[\frac{\partial}{\partial y_{i(calc)}}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,A,B,C}\left(\frac{\partial y_{i(calc)}}{\partial B}\right)_{x_i,A,C}\right] \quad (65)$$

or, substituting equation (63) in equation (65),

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}\right]_{x_i,A,C} = \left[\left(\frac{\partial^2 F}{\partial B \partial y_{i(calc)}}\right)_{x_i,A,C} - \left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i,A,B,C} \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}}\right] \tag{66}$$

Therefore, if we substitute equations (64) and (66) in equation (61), equation (61) is expressible as

$$\left(\frac{\partial^2 Y_i}{\partial A \partial B}\right)_{x_i,y_{i(o)},C} = \left[\frac{\left(\frac{\partial^2 F}{\partial A \partial B}\right)_{x_i,y_{i(calc)},C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i,A,B,C}} - \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}\left(\frac{\partial^2 F}{\partial A \partial y_{i(calc)}}\right)_{x_i,B,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} - \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\left(\frac{\partial^2 F}{\partial B \partial y_{i(calc)}}\right)_{x_i,A,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i,A,B,C}} + \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i,y_{i(calc)},B,C}\left(\frac{\partial F}{\partial B}\right)_{x_i,y_{i(calc)},A,C}\left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i,A,B,C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i,A,B,C}}\right] \tag{67}$$

We can derive expressions for $\left(\frac{\partial^2 Y_i}{\partial A \partial C}\right)_{x_i, y_{i(o)}, B}$ and $\left(\frac{\partial^2 Y_i}{\partial B \partial C}\right)_{x_i, y_{i(o)}, A}$ similar to equation (67). The results are

$$\left(\frac{\partial^2 Y_i}{\partial A \partial C}\right)_{x_i, y_{i(o)}, B} = \left[\frac{\left(\frac{\partial^2 F}{\partial A \partial C}\right)_{x_i, y_{i(calc)}, B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{\left(\frac{\partial F}{\partial C}\right)_{x_i, y_{i(calc)}, A, B} \left(\frac{\partial^2 F}{\partial A \partial y_{i(calc)}}\right)_{x_i, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} - \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C} \left(\frac{\partial^2 F}{\partial C \partial y_{i(calc)}}\right)_{x_i, A, B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} + \frac{\left(\frac{\partial F}{\partial A}\right)_{x_i, y_{i(calc)}, B, C} \left(\frac{\partial F}{\partial C}\right)_{x_i, y_{i(calc)}, A, B} \left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i, A, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i, A, B, C}} \right] \tag{68}$$

$$\left(\frac{\partial^2 Y_i}{\partial B \partial C}\right)_{x_i, y_{i(o)}, A} = \left[\frac{\left(\frac{\partial^2 F}{\partial B \partial C}\right)_{x_i, y_{i(calc)}, A}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)_{x_i, A, B, C}} - \frac{\left(\frac{\partial F}{\partial C}\right)_{x_i, y_{i(calc)}, A, B} \left(\frac{\partial^2 F}{\partial B \partial y_{i(calc)}}\right)_{x_i, A, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} - \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i, y_{i(calc)}, A, C} \left(\frac{\partial^2 F}{\partial C \partial y_{i(calc)}}\right)_{x_i, A, B}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^2_{x_i, A, B, C}} + \frac{\left(\frac{\partial F}{\partial B}\right)_{x_i, y_{i(calc)}, A, C} \left(\frac{\partial F}{\partial C}\right)_{x_i, y_{i(calc)}, A, B} \left(\frac{\partial^2 F}{\partial y^2_{i(calc)}}\right)_{x_i, A, B, C}}{\left(\frac{\partial F}{\partial y_{i(calc)}}\right)^3_{x_i, A, B, C}} \right] \tag{69}$$

The method of obtaining a least squares solution for the constants A, B, and C is as follows.

We have n pairs of the observed quantities, x_i and $y_{i(o)}$.

1. We assume values for A, B, and C.

2. We then calculate the n values of F_i, using equation (2).

3. We next calculate the n values of $y_{i(calc)}$ using an iterative method involving equations (42), (43), and (44).

4. We next calculate the n values of Y_i from equation (4).

5. We next calculate the n values of Y_i^2.

6. Using designated values for $w_{y_{i(o)}}$, we next calculate R, the sum of the weighted squares of the residuals, using equation (5).

7. We next calculate the n values of $(\partial F/\partial y_{i(calc)})$ by differentiation of the analytical expression for F, keeping x_i, A, B, and C constant.

8. We next calculate the n values of $(\partial F/\partial A)$, evaluated at $y_{i(calc)}$, by differentiation of the analytical expression for F, keeping x_i, $y_{i(calc)}$, B, and C constant.

9. We next calculate the n values of $(\partial F/\partial B)$, evaluated at $y_{i(calc)}$, by differentiating the analytical expression for F, keeping x_i, $y_{i(calc)}$, A, and C constant.

10. We next calculate the n values of $(\partial F/\partial C)$, evaluated at $y_{i(calc)}$, by differentiation of the analytical expression for F, keeping x_i, $y_{i(calc)}$, A, and B constant.

11. We next calculate the n values of $(\partial^2 F/\partial A^2)$, evaluated at $y_{i(calc)}$, by differentiating the analytical expression for F twice.

12. We next calculate the n values of $(\partial^2F/\partial B^2)$, evaluated at $y_{i(calc)}$, by differentiation of the analytical expression for F twice.

13. We next calculate the n values of $(\partial^2F/\partial C^2)$, evaluated at $y_{i(calc)}$, by differentiating the analytical expression for F twice.

14. We next calculate the n values of $(\partial^2F/\partial y^2_{i(calc)})$ by differentiating the analytical expression for F twice.

15. We next calculate the n values of $(\partial^2F/\partial A\partial B)$, evaluated at $y_{i(calc)}$, which is obtained from the analytical expression for F.

16. We next calculate the n values of $(\partial^2F/\partial A\partial C)$, evaluated at $y_{i(calc)}$, which is obtained from the analytical expression for F.

17. We next calculate the n values of $(\partial^2F/\partial B\partial C)$, evaluated at $y_{i(calc)}$, which is obtained from the analytical expression for F.

18. We next calculate the n values of $(\partial^2F/\partial y_{i(calc)}\partial A)$ which is obtained from the analytical expression for F.

19. We next calculate the n values of $(\partial^2F/\partial y_{i(calc)}\partial B)$ which is obtained from the analytical expression for F.

20. We next calculate the n values of $(\partial^2F/\partial y_{i(calc)}\partial C)$ which is obtained from the analytical expression for F.

21. We next calculate the n values of $(\partial Y_i/\partial A)$ using equation (49).

22. We next calculate the n values of $(\partial Y_i/\partial B)$ using equation (50).

23. We next calculate the n values of $(\partial Y_i/\partial C)$ using equation (51).

24. We next calculate the n values of $(\partial^2 Y_i/\partial A^2)$ using equation (58).

25. We next calculate the n values of $(\partial^2 Y_i/\partial B^2)$ using equation (59).

26. We next calculate the n values of $(\partial^2 Y_i/\partial C^2)$ using equation (60).

27. We next calculate the n values of $(\partial^2 Y_i/\partial A \partial B)$ using equation (67).

28. We next calculate the n values of $(\partial^2 Y_i/\partial A \partial C)$ using equation (68).

29. We next calculate the n values of $(\partial^2 Y_i/\partial B \partial C)$ using equation (69).

30. We next calculate a_1 from equation (19).

31. We next calculate $a_2 = b_1$ from equation (20).

32. We next calculate $a_3 = c_1$ from equation (21).

33. We next calculate b_2 from equation (22).

34. We next calculate $b_3 = c_2$ from equation (23).

35. We next calculate c_3 from equation (24).

36. We next calculate m_1 from equation (25).

37. We next calculate m_2 from equation (26).

38. We next calculate m_3 from equation (27).

39. We next calculate D_1 from equation (31).

40. We next calculate D_2 from equation (32).

41. We next calculate D_3 from equation (33).

42. We next calculate D_4 from equation (34).

43. We next calculate D_5 from equation (35).

44. We next calculate D_6 from equation (36).

45. We next calculate D_7 from equation (37).

46. We next calculate D_8 from equation (38).

47. We next calculate D_9 from equation (39).

48. We next calculate D_o from equation (40).

49. We next calculate ΔA from equation (28).

50. We next calculate ΔB from equation (29).

51. We next calculate ΔC from equation (30).

52. We now return to step 1 and calculate $A_o + \Delta A$, where A_o is the value originally assumed in step 1.

53. We next calculate $B_o + \Delta B$, where B_o is the value originally assumed in step 1.

54. We next calculate $C_o + \Delta C$, where C_o is the value originally assumed in step 1.

55. Using $(A_o + \Delta A)$, $(B_o + \Delta B)$, and $(C_o + \Delta C)$ as new values of A, B, and C, we proceed to step 2 and repeat steps 2 through 6.

56. At this point, we compare the value of R, the sum of the weighted squares of the residuals, with the initially calculated value of R. If it is smaller, we proceed to step 7 and repeat steps 7 through 55. We continue to repeat the iteration until $m_1 = m_2 = m_3 = 0$ within some predetermined small quantity. Our final values of A, B, and C are our least squares solution for these quantities.

If in step 56 the sum of the weighted squares of the residuals, R, is larger than the initial value of R, the problem is diverging, and at this point, a technique must be used that will lead to convergence. One method that will lead to a smaller sum of R is the negative gradient or the method of steepest descent. The components of a vector in the direction of the negative gradient are given by

$$-\left(\frac{\partial R}{\partial A}\right)^{o};\quad -\left(\frac{\partial R}{\partial B}\right)^{o};\ \text{and}\ -\left(\frac{\partial R}{\partial C}\right)^{o}$$

Thus, if we take

$$\Delta A = -k\left(\frac{\partial R}{\partial A}\right)^{o} \qquad (70)$$

$$\Delta B = -k\left(\frac{\partial R}{\partial B}\right)^{o} \qquad (71)$$

$$\Delta C = -k\left(\frac{\partial R}{\partial C}\right)^{o} \qquad (72)$$

where k is a positive constant, we will move in the direction of the negative gradient; k has been called the size of the step. We determine the size of the step in the following way. We expand R in a Taylor's series expansion, retaining terms through the second derivatives. Thus,

$$R = \left[\begin{array}{l} R^{o} + \left(\frac{\partial R}{\partial A}\right)^{o}\Delta A + \left(\frac{\partial R}{\partial B}\right)^{o}\Delta B + \left(\frac{\partial R}{\partial C}\right)^{o}\Delta C \\ + \frac{1}{2}\left[\left(\frac{\partial^2 R}{\partial A^2}\right)^{o}(\Delta A)^2 + \left(\frac{\partial^2 R}{\partial B^2}\right)^{o}(\Delta B)^2 + \left(\frac{\partial^2 R}{\partial C^2}\right)^{o}(\Delta C)^2\right] \\ + \left(\frac{\partial^2 R}{\partial A\partial B}\right)^{o}\Delta A\Delta B + \left(\frac{\partial^2 R}{\partial A\partial C}\right)^{o}\Delta A\Delta C + \left(\frac{\partial^2 R}{\partial B\partial C}\right)^{o}\Delta B\Delta C \end{array}\right] \tag{73}$$

We now substitute in equation (73) for ΔA, ΔB, and ΔC from equations (70), (71), and (72). The result is

$$R = \left[\begin{array}{l} R^{o} - k\left[\left(\frac{\partial R}{\partial A}\right)^{o^2} + \left(\frac{\partial R}{\partial B}\right)^{o^2} + \left(\frac{\partial R}{\partial C}\right)^{o^2}\right] \\ + \frac{1}{2}k^2\left[\left(\frac{\partial R}{\partial A}\right)^{o^2}\left(\frac{\partial^2 R}{\partial A^2}\right)^{o} + \left(\frac{\partial R}{\partial B}\right)^{o^2}\left(\frac{\partial^2 R}{\partial B^2}\right)^{o} + \left(\frac{\partial R}{\partial C}\right)^{o^2}\left(\frac{\partial^2 R}{\partial C^2}\right)^{o}\right] \\ + k^2\left[\left(\frac{\partial R}{\partial A}\right)^{o}\left(\frac{\partial R}{\partial B}\right)^{o}\left(\frac{\partial^2 R}{\partial A\partial B}\right)^{o} + \left(\frac{\partial R}{\partial A}\right)^{o}\left(\frac{\partial R}{\partial C}\right)^{o}\left(\frac{\partial^2 R}{\partial A\partial C}\right)^{o} + \left(\frac{\partial R}{\partial B}\right)^{o}\left(\frac{\partial R}{\partial C}\right)^{o}\left(\frac{\partial^2 R}{\partial B\partial C}\right)^{o}\right] \end{array}\right] \tag{74}$$

We now differentiate equation (74) with regard to k, set the derivative equal to zero, and solve for k. The result is

$$k = \frac{\left(\frac{\partial R}{\partial A}\right)^{o^2} + \left(\frac{\partial R}{\partial B}\right)^{o^2} + \left(\frac{\partial R}{\partial C}\right)^{o^2}}{\left[\begin{array}{l} \left(\frac{\partial R}{\partial A}\right)^{o^2}\left(\frac{\partial^2 R}{\partial A^2}\right)^{o} + \left(\frac{\partial R}{\partial B}\right)^{o^2}\left(\frac{\partial^2 R}{\partial B^2}\right)^{o} + \left(\frac{\partial R}{\partial C}\right)^{o^2}\left(\frac{\partial^2 R}{\partial C^2}\right)^{o} \\ + 2\left(\frac{\partial R}{\partial A}\right)^{o}\left(\frac{\partial R}{\partial B}\right)^{o}\left(\frac{\partial^2 R}{\partial A\partial B}\right)^{o} + 2\left(\frac{\partial R}{\partial A}\right)^{o}\left(\frac{\partial R}{\partial C}\right)^{o}\left(\frac{\partial^2 R}{\partial A\partial C}\right)^{o} + 2\left(\frac{\partial R}{\partial B}\right)^{o}\left(\frac{\partial R}{\partial C}\right)^{o}\left(\frac{\partial^2 R}{\partial B\partial C}\right)^{o} \end{array}\right]} \tag{75}$$

From equation (5) and equations (19) - (27), it is possible to show that

$$\left(\frac{\partial R}{\partial A}\right)^{o} = -2m_1 \tag{76}$$

$$\left(\frac{\partial R}{\partial B}\right)^{o} = -2m_2 \tag{77}$$

$$\left(\frac{\partial R}{\partial C}\right)^{o} = -2m_3 \tag{78}$$

$$\left(\frac{\partial^2 R}{\partial A^2}\right)^{o} = 2a_1 \tag{79}$$

$$\left(\frac{\partial^2 R}{\partial B^2}\right) = 2b_2 \tag{80}$$

$$\left(\frac{\partial^2 R}{\partial C^2}\right)^{o} = 2c_3 \tag{81}$$

$$\left(\frac{\partial^2 R}{\partial A \partial B}\right)^{o} = 2a_2 = 2b_1 \tag{82}$$

$$\left(\frac{\partial^2 R}{\partial A \partial C}\right)^{o} = 2a_3 = 2c_1 \tag{83}$$

$$\left(\frac{\partial^2 R}{\partial B \partial C}\right)^{o} = 2b_3 = 2c_2 \tag{84}$$

Substituting equations (76) - (84) into equation (75), we have

$$k = \frac{m_1^2 + m_2^2 + m_3^2}{\left[\begin{array}{l} 2a_1m_1^2 + 2b_2m_2^2 + 2c_3m_3^2 \\ + 4a_2m_1m_2 + 4a_3m_1m_3 + 4b_3m_2m_3 \end{array}\right]} \tag{85}$$

We see, therefore, that if the Newton-Raphson method is used in setting up the normal equations, the size of the step to be taken in the direction of the negative gradient can be evaluated very simply from the coefficients appearing in the normal equations, provided the calculation of k leads to a positive quantity. If equation (85) leads to a negative quantity, this means that the curvature of the surface is such that the trial solution is near a maximum and not a minimum. Under these conditions, the negative value of k must be ignored and positive values of k explored on a trial basis.

Of course, the same formal calculation can be made if the Gauss-Newton method is used to set up the normal equations. However, if one is in a region of divergence, this means that the trial solution is far from the true answer. Under these conditions, the residuals will be large and the second summation in the a's, b's, and c's, involving the second derivative terms, will be of importance compared to the first term in the summation. We therefore believe that if it is necessary to use the negative gradient method, it is better to use the Newton-Raphson method in setting up the normal equations and in calculating the size of the step.

If our problem is diverging after the first iteration, we calculate ΔA, ΔB, and ΔC from

$$\Delta A = 2km_1 \tag{86}$$

$$\Delta B = 2km_2 \tag{87}$$

$$\Delta C = 2km_3 \tag{88}$$

with k calculated from equation (85). Using $A_o + \Delta A$, $B_o + \Delta B$, and $C_o + \Delta C$ as new values of our undetermined constants, A, B, and C, we then return to step 2 of the iteration. For each new value of A, B, and C, we solve the normal equations and calculate a new value of R. If we are diverging, we continue with the negative gradient method until the region of convergence is reached. As soon as this happens, we drop the negative gradient method and iterate by solving the normal equations for ΔA, ΔB, and ΔC. A scheme such as this should always lead to convergence.

Although the above scheme should always lead to convergence, the negative gradient method may be tediously slow in entering the region of convergence for the solution of the normal equations. Under these conditions, other schemes can be tried which may enter the region of convergence more rapidly than the negative gradient method.

Some of these methods are: (1) the Hartley (3) method; (2) a modification of the Hartley method due to Strand, Kohl, and Bonham (8); and (3) the method of Marquardt (4).

We do not have enough experience to judge the relative merits of these various methods. In our applications so far, we have not been troubled by lack of convergence.

CALCULATION OF VARIANCES AND COVARIANCES

With the value of our constants determined, the remaining questions to be answered are: (1) What are the variances and covariances of the constants evaluated? (2) What are the variances

of the calculated y_i's and of any other calculated y that reduces F to zero?

To answer these questions, we apply the law for the propagation of errors. This law states that if we have a quantity or function, Q, that is a function of the independent quantities, y_1, y_2, y_3, ... , then the variance of Q is given by

$$S_Q^2 = \sum_{i=1}^{n} \left(\frac{\partial Q}{\partial y_{i(o)}} \right)^2 S_{y_{i(o)}}^2 \tag{89}$$

where S_Q^2 is the variance of Q, and $S_{y_{i(o)}}^2$ is the variance of $y_{i(o)}$.

The value of the constant A that we have evaluated is a function of all of the observed x_i's and of all of the observed y_i's. Since we have assumed there are no random errors in the x_i's, the variances of the x_i's are zero. Then the expression for the variance of A is given by the equation

$$S_A^2 = \sum_{i=1}^{n} \left(\frac{\partial A}{\partial y_{i(o)}} \right)^2 S_{y_{i(o)}}^2 \tag{90}$$

and there will be an equation similar to equation (90) for evaluating the variance of B and of C.

To evaluate equation (90), we must evaluate $(\partial A/\partial y_{i(o)})$ for each $y_{i(o)}$, multiply this quantity by $S_{y_{i(o)}}$, square the product, and then sum the product over all of the observed y_i's.

Now we have a total of n pairs of the observed quantities x_i, y_i. Then our constants to be evaluated are determined by the

solutions of equations (6), (7), and (8). Now suppose we change one of the y_i's, $y_{2(o)}$ say, to $y_{2(o)} + \Delta y_2$. Then on solving equations (6), (7), and (8), we would get new values of $(A + \Delta A)$, $(B + \Delta B)$, and $(C + \Delta C)$ for our constants. Then we can calculate

$$\frac{\partial A}{\partial y_{2(o)}} = \frac{\Delta A}{\Delta y_2} \tag{91}$$

$$\frac{\partial B}{\partial y_{2(o)}} = \frac{\Delta B}{\Delta y_2} \tag{92}$$

$$\frac{\partial C}{\partial y_{2(o)}} = \frac{\Delta C}{\Delta y_2} \tag{93}$$

This means that when $y_{2(o)}$ is changed by a small amount, equations (6), (7), and (8) must still hold exactly. Mathematically, this means that

$$\frac{\partial}{\partial y_{i(o)}} \left(\frac{\partial R}{\partial A}\right)_{x_i, y_{i(o)}, B, C} = 0 \tag{94}$$

$$\frac{\partial}{\partial y_{i(o)}} \left(\frac{\partial R}{\partial B}\right)_{x_i, y_{i(o)}, A, C} = 0 \tag{95}$$

$$\frac{\partial}{\partial y_{i(o)}} \left(\frac{\partial R}{\partial C}\right)_{x_i, y_{i(o)}, A, B} = 0 \tag{96}$$

All of the Y_i's are explicit functions of the $y_{i(o)}$'s, the x_i's, and the constants A, B, and C. In the application of equations (94),

(95), and (96), the derivatives $(\partial R/\partial A)_{x_i, y_{i(o)}, B, C}$, $(\partial R/\partial B)_{x_i, y_{i(o)}, A, C}$, and $(\partial R/\partial C)_{x_i, y_{i(o)}, A, B}$ are to be considered functions of all of the $y_{i(o)}$'s and of the constants evaluated, with A, B, and C being functions of all of the $y_{i(o)}$'s. Under these conditions,

$$\left[\frac{\partial}{\partial y_{i(o)}}\left(\frac{\partial R}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{j\neq i}} = \left[\begin{aligned} &\left(\frac{\partial^2 R}{\partial A^2}\right)_{x_i, y_{i(o)}, B, C}\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left(\frac{\partial^2 R}{\partial B \partial A}\right)_{x_i, y_{i(o)}, C}\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left(\frac{\partial^2 R}{\partial C \partial A}\right)_{x_i, y_{i(o)}, B}\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left[\frac{\partial\left(\frac{\partial R}{\partial A}\right)_{x_i, y_{i(o)}, B, C}}{\partial y_{i(o)}}\right]_{x_i, y_{j\neq i}, A, B, C} \end{aligned}\right] = 0 \qquad (97)$$

$$\left[\frac{\partial}{\partial y_{i(o)}}\left(\frac{\partial R}{\partial B}\right)_{x_i, y_{i(o)}, A, C}\right]_{x_i, y_{j\neq i}} = \left[\begin{aligned} &\left(\frac{\partial^2 R}{\partial B^2}\right)_{x_i, y_{i(o)}, A, C}\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left(\frac{\partial^2 R}{\partial A \partial B}\right)_{x_i, y_{i(o)}, C}\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left(\frac{\partial^2 R}{\partial C \partial B}\right)_{x_i, y_{i(o)}, A}\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ &+ \left[\frac{\partial\left(\frac{\partial R}{\partial B}\right)_{x_i, y_{i(o)}, A, C}}{\partial y_{i(o)}}\right]_{x_i, y_{j\neq i}, A, B, C} \end{aligned}\right] = 0 \qquad (98)$$

$$\left[\frac{\partial}{\partial y_{i(o)}}\left(\frac{\partial R}{\partial C}\right)_{x_i,y_{i(o)},A,B}\right]_{x_i,y_{j\neq i}} = \left[\begin{array}{l} \left(\frac{\partial^2 R}{\partial C^2}\right)_{x_i,y_{i(o)},A,B}\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ + \left(\frac{\partial^2 R}{\partial A \partial C}\right)_{x_i,y_{i(o)},B}\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ + \left(\frac{\partial^2 R}{\partial B \partial C}\right)_{x_i,y_{i(o)},A}\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \\ + \left[\frac{\partial\left(\frac{\partial R}{\partial C}\right)_{x_i,y_{i(o)},A,B}}{\partial y_{i(o)}}\right]_{x_i,y_{j\neq i},A,B,C} \end{array}\right] = 0 \quad (99)$$

Equations (97), (98), and (99) contain terms involving the derivative of each constant with respect to $y_{i(o)}$ and can be solved by elementary algebra for the derivatives of the constants with respect to $y_{i(o)}$. These derivatives are then to be multiplied by $S_{y_{i(o)}}$, the product squared, and then the squared product is to be summed over all of the observed $y_{i(o)}$'s. These sums give us the variance of each constant evaluated.

Differentiating R, given by equation (5), by the constants A, B, and C keeping x_i and $y_{i(o)}$ constant, we have

$$\left(\frac{\partial R}{\partial A}\right)_{x_i,y_{i(o)},B,C} = 2\sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \quad (100)$$

$$\left(\frac{\partial R}{\partial B}\right)_{x_i, y_{i(o)}, A, C} = 2\sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \tag{101}$$

$$\left(\frac{\partial R}{\partial C}\right)_{x_i, y_{i(o)}, A, B} = 2\sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \tag{102}$$

Differentiating equation (100) with regard to each constant, we have

$$\left(\frac{\partial^2 R}{\partial A^2}\right)_{x_i, y_{i(o)}, B, C} = 2\sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i, y_{i(o)}, B, C} + Y_i \left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i, y_{i(o)}, B, C}\right] \tag{103}$$

$$\left[\frac{\partial}{\partial B}\left(\frac{\partial R}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{i(o)}, A, C} = \left[\begin{aligned} &2\sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \\ &+ 2\sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left[\frac{\partial}{\partial B}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{i(o)}, A, C} \end{aligned}\right] \tag{104}$$

$$\left[\frac{\partial}{\partial C}\left(\frac{\partial R}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{i(o)}, A, B} = \left[\begin{aligned} &2\sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \\ &+ 2\sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left[\frac{\partial}{\partial C}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\right]_{x_i, y_{i(o)}, A, B} \end{aligned}\right] \tag{105}$$

Differentiating equation (100) with regard to a single $y_{i(o)}$, keeping A, B, and C constant, we have

$$\left[\frac{\partial\left(\frac{\partial R}{\partial A}\right)_{x_i,y_{i(o)},B,C}}{\partial y_{i(o)}}\right]_{x_i,y_{j\neq i},A,B,C} = 2\left[\begin{array}{l} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial y_{i(o)}}\right)_{x_i,y_{j\neq i},A,B,C}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \\ + w_{y_{i(o)}}Y_i\left[\frac{\partial}{\partial y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\right]_{x_i,y_{j\neq i},A,B,C} \end{array}\right] \qquad (106)$$

The right-hand side of equation (106) reduces to a single term, since a single $y_{i(o)}$ only appears in one term in the summation.

Now from equation (4)

$$Y_i = y_{i(o)} - y_{i(calc)} \qquad (4)$$

so that

$$\left(\frac{\partial Y_i}{\partial y_{i(o)}}\right)_{x_i,y_{j\neq i},A,B,C} = 1 \qquad (107)$$

and it follows that

$$\left[\frac{\partial}{\partial A}\left(\frac{\partial Y_i}{\partial y_{i(o)}}\right)_{x_i,y_{j\neq i},A,B,C}\right]_{x_i,y_{i(o)},B,C} = 0 \qquad (108)$$

Substituting equations (107) and (108) into equation (106), we have

$$\left[\frac{\partial\left(\frac{\partial R}{\partial A}\right)_{x_i,y_{i(o)},B,C}}{\partial y_{i(o)}}\right]_{x_i,y_{j\neq i},A,B,C} = 2w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \quad (109)$$

Substituting equations (103), (104), (105), and (109) in equation (97), we get

$$\left[\begin{aligned} &\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} \sum_{i=1}^{n} w_{y_{i(o)}}\left[\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} + Y_i\left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i,y_{i(o)},B,C}\right] \\ &+ \left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}}\left[\begin{aligned} &\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \\ &+\sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left[\frac{\partial}{\partial B}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\right]_{x_i,y_{i(o)},A,C}\end{aligned}\right] \\ &+ \left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}}\left[\begin{aligned} &\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \\ &+\sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left[\frac{\partial}{\partial C}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\right]_{x_i,y_{i(o)},A,B}\end{aligned}\right] \\ &+ w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \end{aligned}\right] = 0 \quad (110)$$

Upon differentiating equations (101) and (102) with regard to a single $y_{i(o)}$, we obtain two expressions similar to equation (110). We will write equation (110) and the two other equations as

$$a_1\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + b_1\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + c_1\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} = n_1 \quad (111)$$

$$a_2\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + b_2\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + c_2\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} = n_2 \quad (112)$$

$$a_3\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + b_3\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i}} + c_3\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i}} = n_3 \quad (113)$$

where

$$a_1 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} + Y_i\left(\frac{\partial^2 Y_i}{\partial A^2}\right)_{x_i,y_{i(o)},B,C}\right] \quad (114)$$

$$a_2 = b_1 = \left[\begin{array}{l} \displaystyle\sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \\ \displaystyle + \sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left(\frac{\partial^2 Y_i}{\partial B \partial A}\right)_{x_i,y_{i(o)},C} \end{array}\right] \quad (115)$$

$$a_3 = c_1 = \left[\begin{array}{l} \displaystyle\sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C} \\ \displaystyle + \sum_{i=1}^{n} w_{y_{i(o)}} Y_i\left(\frac{\partial^2 Y_i}{\partial C \partial A}\right)_{x_i,y_{i(o)},B} \end{array}\right] \quad (116)$$

$$b_2 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i, y_{i(o)}, A, C} + Y_i \left(\frac{\partial^2 Y_i}{\partial B^2}\right)_{x_i, y_{i(o)}, A, C} \right] \tag{117}$$

$$b_3 = c_2 = \left[\begin{array}{l} \displaystyle\sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \\ \displaystyle + \sum_{i=1}^{n} w_{y_{i(o)}} Y_i \left(\frac{\partial^2 Y_i}{\partial C \partial B}\right)_{x_i, y_{i(o)}, A} \end{array} \right] \tag{118}$$

$$c_3 = \sum_{i=1}^{n} w_{y_{i(o)}} \left[\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i, y_{i(o)}, A, B} + Y_i \left(\frac{\partial^2 Y_i}{\partial C^2}\right)_{x_i, y_{i(o)}, A, B} \right] \tag{119}$$

$$n_1 = -w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \tag{120}$$

$$n_2 = -w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \tag{121}$$

$$n_3 = -w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \tag{122}$$

Notice that the a's, b's, and c's defined by equations (114) - (119) are of the same form as the a's, b's, and c's defined by

equations (19) - (24) which appear in our normal equations. The difference between these two definitions is that equations (114) - (119) apply when the least squares solution is obtained while equations (19) - (24) apply to trial solutions. The a's, b's, and c's given by equations (114) - (119) can be considered as the values of the coefficients in the normal equations for the final iteration.

Solving equations (111), (112), and (113) simultaneously for $(\partial A/\partial y_{i(o)})$, $(\partial B/\partial y_{i(o)})$, and $(\partial C/\partial y_{i(o)})$, we get

$$D_o\left(\frac{\partial A}{\partial y_{i(o)}}\right) = D_1 n_1 + D_2 n_2 + D_3 n_3 \tag{123}$$

$$D_o\left(\frac{\partial B}{\partial y_{i(o)}}\right) = D_4 n_1 + D_5 n_2 + D_6 n_3 \tag{124}$$

$$D_o\left(\frac{\partial C}{\partial y_{i(o)}}\right) = D_7 n_1 + D_8 n_2 + D_9 n_3 \tag{125}$$

where

$$D_1 = b_2 c_3 - b_3 c_2 \tag{126}$$

$$D_2 = b_3 c_1 - b_1 c_3 \tag{127}$$

$$D_3 = b_1 c_2 - b_2 c_1 \tag{128}$$

$$D_4 = a_3 c_2 - a_2 c_3 \tag{129}$$

$$D_5 = a_1 c_3 - a_3 c_1 \tag{130}$$

$$D_6 = a_2c_1 - a_1c_2 \tag{131}$$

$$D_7 = a_2b_3 - a_3b_2 \tag{132}$$

$$D_8 = a_3b_1 - a_1b_3 \tag{133}$$

$$D_9 = a_1b_2 - a_2b_1 \tag{134}$$

$$D_o = a_1D_1 + a_2D_2 + a_3D_3 \tag{135}$$

Notice that the D's defined by equations (126) - (135) are of the same form as the D's defined by equations (31) - (40) and, for the last iteration, when the least squares solution is obtained, they will be identical.

Squaring equation (123) and multiplying by $S^2_{y_{i(o)}}$, we have

$$D_o^2\left(\frac{\partial A}{\partial y_{i(o)}}\right)^2 S^2_{y_{i(o)}} = \left[\begin{array}{l} D_1^2n_1^2S^2_{y_{i(o)}} + D_2^2n_2^2S^2_{y_{i(o)}} + D_3^2n_3^2S^2_{y_{i(o)}} \\ + 2D_1D_2n_1n_2S^2_{y_{i(o)}} + 2D_1D_3n_1n_3S^2_{y_{i(o)}} \\ + 2D_2D_3n_2n_3S^2_{y_{i(o)}} \end{array}\right] \tag{136}$$

But from equation (120)

$$n_1^2S^2_{y_{i(o)}} = w^2_{y_{i(o)}}\left[\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C}\right]S^2_{y_{i(o)}} \tag{137}$$

Now

$$s^2_{y_{i(o)}} = \frac{L^2}{w_{y_{i(o)}}} \tag{138}$$

where L is a constant. Substituting equation (138) in equation (137), we find that

$$n_1^2 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i, y_{i(o)}, B, C} \tag{139}$$

Similarly, we find that

$$n_2^2 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i, y_{i(o)}, A, C} \tag{140}$$

$$n_3^2 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i, y_{i(o)}, A, B} \tag{141}$$

$$n_1 n_2 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \tag{142}$$

$$n_1 n_3 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \tag{143}$$

$$n_2 n_3 s^2_{y_{i(o)}} = L^2 \, w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \tag{144}$$

Substituting equations (139), (140), (141), (142), (143), and (144) into equation (136) and then summing over all of the observed $y_{i(o)}$'s, we have

$$\sum_{i=1}^{n} \left(\frac{\partial A}{\partial y_{i(o)}}\right)^2 s^2_{y_{i(o)}} = s_A^2 = \frac{L^2}{D_o^2} \left[\begin{aligned} & D_1^2 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i, y_{i(o)}, B, C} \\ & + D_2^2 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i, y_{i(o)}, A, C} \\ & + D_3^2 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i, y_{i(o)}, A, B} \\ & + 2D_1D_2 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \\ & + 2D_1D_3 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \\ & + 2D_2D_3 \sum_{i=1}^{n} w_{y_{i(o)}} \left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \end{aligned} \right] \tag{145}$$

Squaring equation (124), multiplying by $S^2_{y_{i(o)}}$ and then summing, we find that

$$\sum_{i=1}^{n}\left(\frac{\partial B}{\partial y_{i(o)}}\right)^2 S^2_{y_{i(o)}} = S_B^2 = \frac{L^2}{D_o^2}\left[\begin{aligned} &D_4^2\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} \\ &+ D_5^2\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i,y_{i(o)},A,C} \\ &+ D_6^2\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i,y_{i(o)},A,B} \\ &+ 2D_4D_5\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} \\ &+ 2D_4D_6\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \\ &+ 2D_5D_6\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \end{aligned}\right] \tag{146}$$

Similarly, from equation (125), we find that

$$\sum_{i=1}^{n}\left(\frac{\partial C}{\partial y_{i(o)}}\right)^2 s^2_{y_{i(o)}} = s^2_C = \frac{L^2}{D^2_o}\left[\begin{array}{l} D^2_7 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} \\ + D^2_8 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i,y_{i(o)},A,C} \\ + D^2_9 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i,y_{i(o)},A,B} \\ + 2D_7D_8 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} \\ + 2D_7D_9 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \\ + 2D_8D_9 \sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \end{array}\right] \tag{147}$$

For the covariances, we find from equations (123), (124), and (125)

$$\sum_{i=1}^{n}\left(\frac{\partial A}{\partial y_{i(o)}}\right)\left(\frac{\partial B}{\partial y_{i(o)}}\right)S^2_{y_{i(o)}} = S^2_{AB} = \frac{L^2}{D_o^2}\left[\begin{array}{l} D_1D_4\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} \\ + D_2D_5\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i,y_{i(o)},A,C} \\ + D_3D_6\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i,y_{i(o)},A,B} \\ +(D_2D_4+D_1D_5)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} \\ +(D_1D_6+D_3D_4)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \\ + (D_2D_6+ D_3D_5)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \end{array}\right] \qquad (148)$$

$$\sum_{i=1}^{n}\left(\frac{\partial A}{\partial y_{i(o)}}\right)\left(\frac{\partial C}{\partial y_{i(o)}}\right)S^2_{y_{i(o)}} = S^2_{AC} = \frac{L^2}{D_o^2}\left[\begin{aligned}
&D_1D_7\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i,y_{i(o)},B,C} \\
&+ D_2D_8\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i,y_{i(o)},A,C} \\
&+ D_3D_9\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i,y_{i(o)},A,B} \\
&+(D_1D_8+D_2D_7)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C} \\
&+(D_1D_9+ D_3D_7)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i,y_{i(o)},B,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B} \\
&+(D_2D_9+ D_3D_8)\sum_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i,y_{i(o)},A,C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i,y_{i(o)},A,B}
\end{aligned}\right] \tag{149}$$

and finally,

$$\sum_{i=1}^{n}\left(\frac{\partial B}{\partial y_{i(o)}}\right)\left(\frac{\partial C}{\partial y_{i(o)}}\right)S^2_{y_{i(o)}} = S^2_{BC} = \frac{L^2}{D_o^2}\left[\begin{array}{l} D_4D_7 \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)^2_{x_i, y_{i(o)}, B, C} \\ + D_5D_8 \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)^2_{x_i, y_{i(o)}, A, C} \\ + D_6D_9 \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial C}\right)^2_{x_i, y_{i(o)}, A, B} \\ +(D_4D_8 + D_5D_7) \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C} \\ +(D_4D_9 + D_6D_7) \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial A}\right)_{x_i, y_{i(o)}, B, C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \\ +(D_5D_9 + D_6D_8) \sum\limits_{i=1}^{n} w_{y_{i(o)}}\left(\frac{\partial Y_i}{\partial B}\right)_{x_i, y_{i(o)}, A, C}\left(\frac{\partial Y_i}{\partial C}\right)_{x_i, y_{i(o)}, A, B} \end{array}\right] \tag{150}$$

Equations (145), (146), (147), (148), (149), and (150) allow us to calculate the variances and covariances of the constants evaluated.

Now D_o is the value of the determinate of the coefficients that appear in the final normal equations. In a linear problem, it is found that D_o may be factored from the right-hand side of equations (145) - (150) so that the expressions for the variances and covariances only involve D_o to the first power in the denominator. This is not true for a non-linear problem as can be shown from the exact derivation given above. The present method in use in calculating variances and covariances is therefore only an approximation. How good this approximation is can only be decided by testing it against the mathematically exact calculation given above. Since this is true, we prefer to calculate variances and covariances using the mathematically exact equations given above.

We now answer the question, how do we calculate the variance of a calculated y which reduces the function F to zero for a given x? The answer to this question is obtained as follows: y is a function of x, and through the constants evaluated, is a function of all of the observed $y_{i(o)}$'s and x_i's. When we apply the law for the propagation of errors, we have

$$S_y^2 = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial y_{i(o)}}\right)^2_{y_{j\neq i},x,x_i} S^2_{y_{i(o)}} \qquad (151)$$

We calculate $\left(\frac{\partial y}{\partial y_{i(o)}}\right)_{y_{j\neq i},x,x_i}$ from equation (1). Differentiating equation (1), we have

$$\left[\begin{aligned}&\left(\frac{\partial F}{\partial A}\right)_{x,y,B,C}\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}+\left(\frac{\partial F}{\partial B}\right)_{x,y,A,C}\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}\\&+\left(\frac{\partial F}{\partial C}\right)_{x,y,A,B}\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}+\left(\frac{\partial F}{\partial y}\right)_{x,A,B,C}\left(\frac{\partial y}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i,x}\end{aligned}\right]=0 \qquad (152)$$

Solving equation (152) for $\left(\frac{\partial y}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i,x}$, we get

$$\left(\frac{\partial y}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i,x}=-\frac{1}{\left(\frac{\partial F}{\partial y}\right)_{x,A,B,C}}\left[\begin{aligned}&\left(\frac{\partial F}{\partial A}\right)_{x,y,B,C}\left(\frac{\partial A}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}\\&+\left(\frac{\partial F}{\partial B}\right)_{x,y,A,C}\left(\frac{\partial B}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}\\&+\left(\frac{\partial F}{\partial C}\right)_{x,y,A,B}\left(\frac{\partial C}{\partial y_{i(o)}}\right)_{y_{j\neq i},x_i}\end{aligned}\right] \qquad (153)$$

Squaring equation (153), multiplying by $S^2_{y_{i(o)}}$, and then summing over all the observed $y_{i(o)}$'s, we obtain

$$S_y^2 = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial y_{i(o)}}\right)^2_{y_{j\neq i},x,x_i} S^2_{y_{i(o)}} = \frac{1}{\left(\frac{\partial F}{\partial y}\right)^2_{x,A,B,C}} \left[S_A^2\left(\frac{\partial F}{\partial A}\right)^2_{x,y,B,C} + S_B^2\left(\frac{\partial F}{\partial B}\right)^2_{x,y,A,C} + S_C^2\left(\frac{\partial F}{\partial C}\right)^2_{x,y,A,B} + 2S^2_{AB}\left(\frac{\partial F}{\partial A}\right)_{x,y,B,C}\left(\frac{\partial F}{\partial B}\right)_{x,y,A,C} + 2S^2_{AC}\left(\frac{\partial F}{\partial A}\right)_{x,y,B,C}\left(\frac{\partial F}{\partial C}\right)_{x,y,A,B} + 2S^2_{BC}\left(\frac{\partial F}{\partial B}\right)_{x,y,A,C}\left(\frac{\partial F}{\partial C}\right)_{x,y,A,B} \right] \tag{154}$$

where the variances and covariances in equation (154) are to be calculated from equations (145), (146), (147), (148), (149), and (150). Equation (154) is the general formula for calculating the variance of a calculated y, regardless of the functional relationship between y and x and the parameters, A, B, and C.

This concludes the derivations of the formulas used by us in the solution of general non-linear least squares problems.

In a later paper, we will apply these general formulas to the reduction of PVT data obtained by the Burnett method.

REFERENCES

1. Birge, Raymond T. The Calculation of Errors by the Method of Least Squares. Phys. Rev., v. 40, Apr. 15, 1932, pp. 207-227.

2. Burnett, E. S. Compressibility Determinations Without Volume Measurements. J. Appl. Mech., v. 3A, December 1936, pp. A136-A140.

3. Hartley, H. O. The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares. Technometrics, v. 3, No. 2, May 1961, pp. 269-280.

4. Marquardt, Donald W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. of the Soc. for Ind. and Appl. Math., v. 11, No. 2, June 1963, pp. 431-441.

5. Merriman, Mansfield. A Textbook on the Method of Least Squares. John Wiley and Sons, Inc., New York, N. Y., 8th ed., 1911, pp. 75-79.

6. Moore, Roger Hughes. On the Least Squares Estimation of Parameters in Nonlinear Models. Ph. D. thesis, Oklahoma State University, Stillwater, Oklahoma, 1962. University Microfilms, Inc., Ann Arbor, Michigan, 63-4067.

7. Scarborough, James B. Numerical Mathematical Analysis. The John Hopkins Press, Baltimore, Maryland, 4th ed., 1958, pp. 192-198.

8. Strand, T. G., D. A. Kohl, and R. A. Bonham. Modification to the Newton-Raphson Method for the Fitting of Nonlinear Functions by Least Squares. J. Chem. Phys., v. 39, No. 5, Sept. 1, 1963, pp. 1307-1310.